

Computative Graphematik – 34 Jahre später

Hanna Lüscho

(Institut für Germanistik, Carl von Ossietzky Universität Oldenburg)

Wie gut hält die Methodik in der historischen Graphematik dem Zahn der Zeit stand, besonders, wenn Computer mit im Spiel sind? Wo liegen die Probleme und Möglichkeiten, die uns die digitale Datenverarbeitung bietet? Sind die Herangehensweisen der 60er bis 80er Jahre heute noch aktuell, oder können sie nur noch als primitiv abgestempelt werden?

Zur Beantwortung dieser Fragen sollen exemplarisch das Buch ‚*Computative Graphematik. Dokumentation eines Programmsystems zur schreibsprachlichen Analyse historischer Texte*‘ von Gerhard Koller (1985) und das darin beschriebene Vorgehen aus heutiger Perspektive – dreieinhalb Dekaden später – kritisch beäugt und möglichst originalgetreu reproduziert werden.

Einige der damaligen Einschränkungen sind heute in der Regel nicht mehr der Rede wert (z.B.: Wie codiert man Umlaute und Großbuchstaben? Für wie viele Wortformen reicht der Speicherplatz? Wie groß muss eine Stichprobe sein, um verlässliche Aussagen über den Text zuzulassen?), andere sind hingegen weiterhin sehr relevant (z.B.: Für wie viele Wortformen reicht der Arbeitsspeicher? Kann ich ein 100 Millionen Tokens großes Korpus problemlos an meinem PC bearbeiten?).

Natural Language Processing ist heute in aller Munde, hierbei aber weniger in der Graphematik als in anderen Teildisziplinen der Linguistik angelangt; meist beginnt die automatische Verarbeitung erst auf der Ebene der Tokenisierung. Auf der Zeichenebene sind daher auch weiterhin ‚rustikalere‘ Vorgehensweisen denkbar.

Neben einer Veranschaulichung von Interdisziplinarität soll außerdem dafür plädiert werden, sich mehr über Methoden zu unterhalten und dabei auch ruhig konkreter zu werden.

Literatur

Koller, Gerhard (1985): *Computative Graphematik. Dokumentation eines Programmsystems zur schreibsprachlichen Analyse historischer Texte. Zeitschrift für Dialektologie und Linguistik*, Beihefte, Heft 48. Stuttgart: Franz Steiner Verlag.